PERGAMON

Applied Mathematics Letters 15 (2002) 197–202

Applied Mathematics Letters

www.elsevier.com/locate/aml

Finite Element Analysis for a Nonlinear Diffusion Model in Image Processing

JICHUN LI
Dept. of Mathematical Sciences, University of Nevada, Las Vegas
4505 Maryland Parkway, Box 454020, Las Vegas, NV 89154-4020, U.S.A.

(*Received November 2000; revised and accepted March 2001*)

Abstract—The optimal error estimate $O(h^{k+1})$ for a popular nonlinear diffusion model widely used in image processing is proved for the standard k^{th}-order ($k \geq 1$) conforming tensor-product finite elements in the L^2-norm. The optimal L^2-estimate is obtained by the integral identity technique [1–3] without using the classic Nitsche duality argument [4].

Keywords—Finite element analysis, Nonlinear diffusion, Image processing.

1. INTRODUCTION

Since Perona and Malik [5] introduced a nonlinear diffusion model for image denoising and segmentation in 1987, many other nonlinear diffusion models have been proposed and have proved to be very powerful in the processing of 2-D and 3-D images [6–13]. Also there exist many nonlinear reaction-diffusion systems applied to image restoration, texture generation, and halftoning. More details about the applications of the nonlinear diffusion in image processing can be found in [14], which includes over 400 references. How to efficiently solve these highly nonlinear diffusion equations is very challenging. Compared to the extensively discussed finite difference schemes, very few finite element methods (FEMs) have been investigated [15,16], especially the theoretical convergence analysis for FEM is missing, which motivates this work.

For simplicity, we only consider the following popular model [6,15,16] proposed by Catté, Lions, Morel and Coll [6]:

$$\frac{\partial u}{\partial t} - \nabla \cdot (g(|\nabla G_\sigma * u|)\nabla u) = 0, \qquad (x,t) \in \Omega \times I, \tag{1}$$

$$\frac{\partial u}{\partial \mu} = 0, \qquad (x,t) \in \partial\Omega \times I, \tag{2}$$

$$u(x,0) = u_0, \qquad x \in \Omega, \tag{3}$$

The author would like to thank the referees for their valuable comments and suggestions.

0893-9659/02/$ - see front matter Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX
PII: S0893-9659(01)00118-5

where $I = (0,T)$, $\Omega = (0,1)^2$, and u_0 is the degraded image to be enhanced and smoothed. Here $g : R^+ \to R^+$ is a decreasing function with $g(0) = 1$, $\lim_{s\to\infty} g(s) = 0$, and $g(\sqrt{s})$ is smooth, e.g., $g(s) = (1+s^2)^{-1}$. G_σ is a smoothing kernel such as the Gaussian filter

$$G_\sigma(x) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{|x|^2}{4\sigma}\right), \qquad x = (x_1, x_2) \in R^2,$$

where the constant $\sigma > 0$. Let us note that [6, p. 187]

$$|\nabla G_\sigma * u| = \left[\sum_{i=1}^{2} \left(\frac{\partial G_\sigma}{\partial x_i} * \tilde{u}\right)^2\right]^{1/2},$$

where $*$ denotes the standard convolution and $\tilde{u}$ is an extension of u to R^2 [6, p. 187].

With all the above assumptions, Catté *et al.* proved that [6, Theorem 2.1]: for $u_0 \in L^2(\Omega)$, there exists a unique solution $u \in C^\infty(\overline{\Omega} \times I)$ for (1)–(3). Furthermore, for any $w \in L^\infty(I; L^2(\Omega))$, there exists a constant $\nu > 0$ such that [6, p. 188]

$$g(|\nabla G_\sigma * w|) \geq \nu, \qquad \text{a.e. in } \Omega \times I, \tag{4}$$

where ν depends only on g, G, and $\|u_0\|_{L^2(\Omega)}$. Here

$$L^\infty\left(I; L^2(\Omega)\right) = \left\{w : [0,T] \to L^2(\Omega), \|w\|_{L^\infty(I;L^2(\Omega))} = \max_{0\leq t\leq T} \|w\|_{L^2(\Omega)}(t) < \infty\right\}.$$

The weak formulation of (1)–(3) is given by finding $u \in L^2(I, H^1(\Omega))$ with $u(x,0) = u_0$ such that

$$\left(\frac{\partial u}{\partial t}, v\right) + (g(|\nabla G_\sigma * u|)\nabla u, \nabla v) = 0, \qquad \forall\, v \in H^1(\Omega), \tag{5}$$

where $H^k(\Omega)$ denotes the standard Sobolev space [4] with the norm $\|\cdot\|_{H^k(\Omega)}$.

The FEM for approximating the solution u of (1)–(3) is: find $u_h \in S_h$ such that

$$\left(\frac{\partial u_h}{\partial t}, v\right) + (g(|\nabla G_\sigma * u_h|)\nabla u_h, \nabla v) = 0, \qquad \forall\, v \in S_h, \tag{6}$$

with the initial approximation $u_h^0 = u_h(x,0) \in S_h$ as the standard L^2-projection

$$\left(u_h^0 - u_0, v\right) = 0, \qquad \forall\, v \in S_h, \tag{7}$$

where $S_h \subseteq H^1(\Omega)$ is the standard k^{th} ($k \geq 1$) order conforming tensor-product finite elements Q_k on an arbitrary quasi-uniform rectangular partitions T_h of Ω of maximum mesh size h.

We like to remark that usually an image is modelled as a real valued function, representing the values of the graylevel intensity, defined in some rectangular domain. Hence, the tensor-product finite elements become the natural choice. For example, Preußer and Rumpf [16] used bilinear and trilinear finite elements in their modelling.

In this paper, by using the integral identity technique [1–3], we will prove the following optimal error estimate in the L^2-norm.

THEOREM 1.1. *Let u be the solution of (1)–(3) and u_h be the FEM solution of (6),(7), then we have*

$$\|u - u_h\|_{L^\infty(I;L^2(\Omega))} \leq Ch^{k+1}. \tag{8}$$

The organization of this paper is as follows. In Section 2, the stability and uniqueness of our FEM are presented. Section 3 is devoted to the proof of Theorem of 1.1. Some conclusions are given in the last section.

Throughout the paper, C (or C_i) will denote a generic positive constant, which is independent of the mesh size h and may be of different value at each occurrence.

2. PRELIMINARIES

To develop an *a priori* error estimate for the scheme derived above, we need to examine its stability.

LEMMA 2.1. *For the solution u_h of (6),(7), we have*

$$\|u_h\|^2_{L^2(\Omega)}(t) + \int_0^t \|\nabla u_h\|^2_{L^2(\Omega)}(t)\,dt \le C\|u_0\|^2_{L^2(\Omega)}, \qquad \forall t \in [0,T].$$

PROOF. By letting $v = u_h$ in (6), we obtain

$$\frac{1}{2}\frac{d}{dt}\|u_h\|^2_{L^2(\Omega)} + (g(|\nabla G_\sigma * u_h|)\nabla u_h, \nabla u_h) = 0,$$

which along with (4) gives us

$$\frac{1}{2}\frac{d}{dt}\|u_h\|^2_{L^2(\Omega)} + \nu\|\nabla u_h\|^2_{L^2(\Omega)} \le 0. \tag{9}$$

Integrating (9) in time from 0 to t, we have

$$\frac{1}{2}\|u_h\|^2_{L^2(\Omega)}(t) + \nu \int_0^t \|\nabla u_h\|^2_{L^2(\Omega)}(t)\,dt \le \frac{1}{2}\|u_h^0\|^2_{L^2(\Omega)}. \tag{10}$$

On the other hand, letting $v = u_h^0$ in (7) and using Cauchy-Schwarz inequality, we obtain

$$\left\|u_h^0\right\|_{L^2(\Omega)} \le \|u_0\|_{L^2(\Omega)},$$

which along with (10) completes the proof. ∎

To obtain the optimal error estimate in the L^2-norm, we need the following special interpolation operator $i_h^k w : w \in C^0(\bar\Omega) \to Q_k$ defined on each rectangular element τ of Ω by the following conditions [17, p. 108; 3, p. 332]:

$$\begin{aligned}
& i_h^k w(a_i) = w(a_i), && && 1 \le i \le 4,\\
& \int_{l_j} (i_h^k w - w)v = 0, && \forall\, v|_{l_j} \in P_{k-2}(l_j), && 1 \le j \le 4,\\
& \int_\tau (i_h^k w - w)v = 0, && \forall\, v|_\tau \in Q_{k-2}(\tau) &&
\end{aligned} \tag{11}$$

for $k \ge 2$, where a_i and l_j denote for the vertices and edges of τ. Here P_k is the k^{th}-order polynomial in 1-D. Note that when $k = 1$, i_h^k is defined by (11) only, i.e., i_h^1 is the standard bilinear interpolation.

By [17], i_h^k is well defined and satisfies the following property:

$$\left\|i_h^k w - w\right\|_{L^2(\Omega)} \le Ch^{k+1}\|w\|_{H^{k+1}(\Omega)}. \tag{12}$$

LEMMA 2.2. *For any smooth function $a(x)$ and $w \in H^{k+2}(\Omega)$, we have*

$$\left|\left(a(x)\nabla\left(i_h^k w - w\right), \nabla v\right)\right| \le Ch^{k+1}\|w\|_{H^{k+2}(\Omega)}\|\nabla v\|_{L^2(\Omega)}, \qquad \forall v \in S_h.$$

PROOF. The proof follows directly from Lemma 3 of [3] and the corresponding integral identity for the derivative with respect to y:

$$\int_\tau a(x)\frac{\partial(i_h^k w - w)}{\partial y}\frac{\partial v}{\partial y} = O\left(h^{k+1}\right)\|w\|_{H^{k+2}(\tau)}\left\|\frac{\partial v}{\partial y}\right\|_{L^2(\tau)},$$

which can be proved in the same way as Lemma 3 (II) of [3]. ∎

LEMMA 2.3.
$$\left\|u_h^0 - u_0\right\|_{L^2(\Omega)} \le Ch^{k+1}\|u_0\|_{H^{k+1}(\Omega)}.$$

PROOF. By letting $v = u_h^0 - i_h^k u_0 \in S_h$ in (7), we obtain
$$\begin{aligned} 0 &= \left(u_h^0 - u_0, u_h^0 - i_h^k u_0\right) \\ &= \left(u_h^0 - u_0, u_h^0 - u_0\right) + \left(u_h^0 - u_0, u_0 - i_h^k u_0\right), \end{aligned}$$

from which and (12), we have
$$\left\|u_h^0 - u_0\right\|_{L^2(\Omega)}^2 \le \left\|u_h^0 - u_0\right\|_{L^2(\Omega)} \left\|u_0 - i_h^k u_0\right\|_{L^2(\Omega)} \le Ch^{k+1} \left\|u_h^0 - u_0\right\|_{L^2(\Omega)},$$

which completes the proof. ∎

3. PROOF OF THEOREM 1.1

Denote $\eta = i_h^k u - u_h, \delta = i_h^k u - u$. By subtracting (6) from (5) with $v = v_h$, we obtain the error equation
$$\left(\frac{\partial(u-u_h)}{\partial t}, v_h\right) + (g(|\nabla G_\sigma * u|)\nabla(u-u_h) + (g(|\nabla G_\sigma * u|) - g(|\nabla G_\sigma * u_h|))\nabla u_h, \nabla v_h) = 0, \quad \forall\, v_h \in S_h.$$

Then by letting $v_h = \eta$ in the above equation and reorganizing it, we have
$$\left(\frac{\partial \eta}{\partial t}, \eta\right) + (g(|\nabla G_\sigma * u|)\nabla \eta, \nabla \eta) = \left(\frac{\partial \delta}{\partial t}, \eta\right) + T_1 + T_2, \tag{13}$$

where
$$\begin{aligned} T_1 &= (g(|\nabla G_\sigma * u|)\nabla \delta, \nabla \eta), \\ T_2 &= -(g(|\nabla G_\sigma * u|) - g(|\nabla G_\sigma * u_h|))\nabla u_h, \nabla \eta). \end{aligned}$$

By (4), we have
$$(g(|\nabla G_\sigma * u|)\nabla \eta, \nabla \eta) \ge \nu \|\nabla \eta\|_{L^2(\Omega)}^2. \tag{14}$$

By (12) and Cauchy-Schwarz inequality, we obtain
$$\begin{aligned} \left|\left(\frac{\partial \delta}{\partial t}, \eta\right)\right| &\le Ch^{k+1} \left\|\frac{\partial u}{\partial t}\right\|_{H^{k+1}(\Omega)} \|\eta\|_{L^2(\Omega)} \\ &\le Ch^{2(k+1)} \left\|\frac{\partial u}{\partial t}\right\|_{H^{k+1}(\Omega)}^2 + C\|\eta\|_{L^2(\Omega)}^2. \end{aligned} \tag{15}$$

By Lemma 2.2, we have
$$\begin{aligned} |T_1| &\le Ch^{k+1}\|u\|_{H^{k+2}(\Omega)} \|\nabla \eta\|_{L^2(\Omega)} \\ &\le \frac{\nu}{4}\|\nabla \eta\|_{L^2(\Omega)}^2 + \frac{C}{\nu} h^{2(k+1)} \|u\|_{H^{k+2}(\Omega)}^2. \end{aligned} \tag{16}$$

By the property [6, Equation (2.9)] and (12), we obtain
$$\begin{aligned} |T_2| &\le \|g(|\nabla G_\sigma * u|) - g(|\nabla G_\sigma * u_h|)\|_{L^\infty(\Omega)} \|\nabla u_h\|_{L^2(\Omega)} \|\nabla \eta\|_{L^2(\Omega)} \\ &\le C\|u - u_h\|_{L^2(\Omega)} \|\nabla u_h\|_{L^2(\Omega)} \|\nabla \eta\|_{L^2(\Omega)} \\ &\le C\left(\|\eta\|_{L^2(\Omega)} + h^{k+1}\|u\|_{H^{k+1}(\Omega)}\right) \|\nabla u_h\|_{L^2(\Omega)} \|\nabla \eta\|_{L^2(\Omega)} \\ &\le \frac{\nu}{4}\|\nabla \eta\|_{L^2(\Omega)}^2 + \frac{C}{\nu}\left(\|\eta\|_{L^2(\Omega)}^2 + h^{2(k+1)}\|u\|_{H^{k+1}(\Omega)}^2\right) \|\nabla u_h\|_{L^2(\Omega)}^2. \end{aligned} \tag{17}$$

By substituting (14)–(17) into (13), we have

$$\frac{d}{dt}\|\eta\|^2_{L^2(\Omega)} \le C_1 h^{2(k+1)}\left(1+\|\nabla u_h\|^2_{L^2(\Omega)}\right) + C_2\left(1+\|\nabla u_h\|^2_{L^2(\Omega)}\right)\|\eta\|^2_{L^2(\Omega)}, \tag{18}$$

where the positive constants C_1 and C_2 depend on $\|\frac{\partial u}{\partial t}\|_{H^{k+1}(\Omega)}, \|u\|_{H^{k+2}(\Omega)}, \nu$, but independent of h. Multiplying (18) by $e^{-C_2\int_0^t (1+\|\nabla u_h\|^2_{L^2(\Omega)})\,dt}$ on both sides, we obtain

$$\begin{aligned}
&\frac{d}{dt}\left[e^{-C_2\int_0^t\left(1+\|\nabla u_h\|^2_{L^2(\Omega)}\right)dt}\cdot\|\eta\|^2_{L^2(\Omega)}\right]\\
&\le C_1 h^{2(k+1)}\left(1+\|\nabla u_h\|^2_{L^2(\Omega)}\right)e^{-C_2\int_0^t\left(1+\|\nabla u_h\|^2_{L^2(\Omega)}\right)dt}\\
&\le C_1 h^{2(k+1)}\left(1+\|\nabla u_h\|^2_{L^2(\Omega)}\right).
\end{aligned} \tag{19}$$

By integrating (19) in time from 0 to t, and using Lemmas 2.1 and 2.3, we have

$$\begin{aligned}
\|\eta\|^2_{L^2(\Omega)}(t) &\le \left[\|\eta\|^2_{L^2(\Omega)}(0) + C_1 h^{2(k+1)}\int_0^t\left(1+\|\nabla u_h\|^2_{L^2(\Omega)}\right)dt\right]e^{C_2\int_0^t\left(1+\|\nabla u_h\|^2_{L^2(\Omega)}\right)dt}\\
&\le C_3^2 h^{2(k+1)},
\end{aligned} \tag{20}$$

which gives us

$$\|\eta\|_{L^\infty(I;L^2(\Omega))} \le C_3 h^{k+1}, \tag{21}$$

where the constant $C_3 > 0$ depends on C_1, C_2, and T.

By (21), (12), and the triangle inequality, we have

$$\|u-u_h\|_{L^\infty(I;L^2(\Omega))} \le \|\eta\|_{L^\infty(I;L^2(\Omega))} + \|\delta\|_{L^\infty(I;L^2(\Omega))} \le Ch^{k+1},$$

which completes our proof.

5. CONCLUSIONS

The optimal error estimate $O(h^{k+1})$ for a popular nonlinear diffusion model widely used in image processing is proved for the standard k^{th}-order ($k \ge 1$) conforming tensor-product finite elements in the L^2-norm. We like to remark that with our integral identity technique [1–3], the optimal L^2-estimate is obtained without using the classic Nitsche duality argument [4]. Similar results for fully discrete-time Galerkin methods [4,18] can be pursued without much difficulty.

REFERENCES

1. A. Zhou and J. Li, The full approximation accuracy for the Stream-Function-Vorticity-Pressure Method, *Numer. Math.* **68**, 427–435, (1994).
2. Q. Lin and N.N. Yan, *The Construction and Analysis of High Accurate Finite Element Methods,* (in Chinese), Hebei University Press, Hebei, China, (1996).
3. J. Li, Full-order convergence of a mixed finite element method for fourth-order elliptic equations, *J. Math. Anal. Appl.* **230**, 329–349, (1999).
4. G. Fairweather, *Finite Element Galerkin Methods for Differential Equations*, Marcel Dekker, New York, (1978).
5. P. Perona and J. Malik, Scale space and edge detection using anisotropic diffusion, *IEEE Trans. Pattern Anal. Mach. Intell.* **12**, 629–639, (1990); First appeared in *IEEE Computer Society Workshop on Computer Vision*, (1987).
6. F. Catté, P.-L. Lions, J.-M. Morel and T. Coll, Image selective smoothing and edge detection by nonlinear diffusion, *SIAM J. Numer. Anal.* **29**, 182–193, (1992).
7. L. Alvarez, P.-L. Lions and J.-M. Morel, Image selective smoothing and edge detection by nonlinear diffusion II, *SIAM J. Numer. Anal.* **29**, 845–866, (1992).
8. S. Osher and L. Rudin, Feature-oriented image enhancement using shock filters, *SIAM J. Numer. Anal.* **27**, 919–949, (1990).

9. C.B. Price, P. Wambacq and A. Oosterlinck, Computing with reaction-diffusion systems: Applications in image processing, In *Continuation and Bifurcations: Numerical Techniques and Applications*, (Edited by D. Roose *et al.*), pp. 379–387, Kluwer, (1990).
10. T.F. Chan, G.H. Golub and P. Mulet, A nonlinear primal-dual method for total-variation based image restoration, In *ICAOS '96: Images, Wavelets and PDEs*, (Edited by M.O. Berger, R. Deriche, I. Herlin, J. Jaffré and J.-M. Morel), pp. 241–252, Springer, London, (1996).
11. Y. Chen, B.C. Vemuri and L. Wang, Image denoising and segmentation via nonlinear diffusion, *Computers Math. Applic.* **39** (5/6), 131–149, (2000).
12. V. Caselles, J.-K. Morel, G. Sapiro and A. Tannenbaum, Editors, Special issue on partial differential equations and geometry-driven diffusion in image processing and analysis, *IEEE Trans. Image Proc.* **7** (3), (1998).
13. J. Weickert and C. Schnörr, PDE-based preprocessing of medical images, Technical Report 8/2000, Dept. of Mathematics and Computer Science, Univ. of Mannheim, Germany, (February 2000).
14. J. Weickert, *Anisotropic Diffusion in Image Processing*, Teubner-Verlag, Stuttgart, (1998).
15. E. Bänsch and K. Mikula, A coarsening finite element strategy in image selective smoothing, *Computing and Visualization in Science* **1**, 53–61, (1997).
16. T. Preußer and M. Rumpf, An adaptive finite element method for large scale image processing, *J. Visual Communication and Image Representation* **11**, 183–195, (2000).
17. V. Girault and P.A. Raviart, *Finite Element Methods for Navier-Stokes Equations—Theory and Algorithms*, Springer-Verlag, (1986).
18. M.F. Wheeler, *A priori* L^2 error estimates for Galerkin approximations to parabolic partial differential equations, *SIAM J. Numer. Anal.* **10**, 723–759, (1973).